APPENDIX B

Gene, 100 (1991) 251–255

GENE 03961

Custom polymerase-chain-reaction engineering of a plant expression vector

(Recombinant DNA; cauliflower mosaic virus 35S gene-regulatory elements; 5′-untranslated region; viral coat protein-encoding genes; PCR; CPCR)

Jerry L. Slightom

Molecular Biology Unit 7242, The Upjohn Company, Kalamazoo, MI 49007 (U.S.A.)

Received by R. Wu: 17 September 1990
Revised: 10 November 1990
Accepted: 22 November 1990

SUMMARY

Polymerase-chain-reaction (PCR) amplification combined with custom-synthesized oligodeoxyribonucleotide (oligo) primers can be used to make complex genetic engineering steps (e.g., translational fusions) easy. Much of the complexity of the engineering steps can be incorporated into the custom oligo primers. Using this technique, a plant constitutive expression vector, pUC18cpexp, was constructed. This vector is based on the cauliflower mosaic virus 35S gene-regulatory elements and the cucumber mosaic virus coat protein-encoding gene (*cp*) 5′-untranslated region. Use of this vector is demonstrated by modifying the *cp* genes of several plant viruses and cloning them into pUC18cpexp. Because the construction and use of this vector system require custom oligo primer synthesis and PCR amplification, the technique is referred to as custom PCR engineering.

INTRODUCTION

The PCR amplification technique, using the thermostable *Taq* polymerase (Saiki et al., 1988), has many uses in molecular biology procedures, which include molecular cloning, nt sequencing, and sequence modifications (Innis et al., 1989). A unique feature of PCR is its ability to amplify discrete regions of DNA and to place useful restriction enzyme sites adjacent to the amplified region. In this report the engineering of the plant expression vector pUC18cpexp using CPCR amplification is described. This vector contains the regulatory elements of the CaMV 35S gene (Pietrzak et al., 1986) and most of the CMV 5′-untranslated region (Quemada et al., 1989). Several viral 5′-untranslated regions have been shown to be translational enhancers in plant expression systems (Gallie et al., 1987; Jobling and Gehrke, 1987). In this example, the 5′-untranslated region of the CMV *cp* gene was selected because of its availability; it appears to also serve as a translational enhancer (Quemada et al., 1991). An *Nco*I cloning site was placed at the junction of the CMV-translation start codon (ATG) and 35S terminator DNA region. Even though this expression cassette contains only the *Nco*I site for cloning, its use can be greatly extended if the targeted coding regions are also PCR-amplified to add flanking *Nco*I sites. MacFerrin et al. (1990) have reported a similar use of PCR engineering for the construction of *Escherichia coli* expression vectors.

Correspondence to: Dr. J.L. Slightom, Molecular Biology Unit 7242, The Upjohn Company, Kalamazoo, MI 49007 (U.S.A.) Tel. (616)385-5304; Fax (616)385-7373.

Abbreviations: bp, base pair(s); CaMV, cauliflower mosaic virus; CMV, cucumber mosaic virus; CP, coat protein; *cp*, gene (DNA) encoding CP; CPCR, custom PCR; kb, kilobase(s) or 1000 bp; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; PCR, polymerase chain reaction; PRV, papaya ringspot virus; SV40, simian virus 40; WMV, watermelon mosaic virus; ZYMV, zucchini yellow mosaic virus.

EXPERIMENTAL AND DISCUSSION

(a) CPCR construction of pUC18cpexp

The *cp* gene of CMV was engineered for expression in plants by cloning it between the regulatory sequences of the CaMV 35S gene (Quemada et al., 1991). The engineering of the CMV *cp* gene was straightforward because it contains a 5′-untranslated region (70 nt) and a translation start codon (ATG) (Quemada et al., 1989). Thus it was cloned (using convenient restriction enzyme sites) between the regulatory elements of the CaMV 35S gene to obtain the clone pUC1813cpCMV19 (Fig. 1). However, the *cp* genes of the potyvirus group of plant viruses (which include PRV, WMVII and ZYMV) are much more difficult to engineer because their proteins are encoded within a long single polypeptide (see review by Dougherty and Carrington, 1988). The individual proteins are cleaved from the polyprotein by proteases that are part of the polyprotein. Thus, for the expression of a potyvirus *cp* gene in plants both the important transcriptional and translational elements must be added.

The CMV *cp* gene expression vector pUC1813cpCMV19 contains all the regulatory elements needed for plant expression, but it cannot be conveniently used for the expression of potyvirus *cp* genes because each construct would require a different translational fusion. However, by using CPCR amplification of the vector pUC1813cpCMV19 a more useful plant expression vector, pUC18cpexp, was constructed that allows translational fusions to be made at the site of the translation start codon (ATG). The restriction enzyme recognition site for *Nco*I (CCATGG) was selected because it shares considerable identity with the plant translation consensus sequence (AACAATGGC) described by Lutcke et al. (1987).

CPCR amplifications of the regulatory elements of pUC1813cpCMV19 were done using the oligo primer sets shown in Fig. 1. The oligo primers JLS-81 and -82 (Fig. 1) amplified a 400-bp region that contains 330 bp of the CaMV 35S promoter and the 70-bp CMV 5′-untranslated region. The oligo primers JLS-83 and -84 (Fig. 1) amplified a 200-bp region that contains CaMV 35S terminator region, including the poly(A)-addition signal. These oligo primers added *Hind*III sites to the 5′ and 3′ ends of these CaMV 35S promoter and terminator fragments and *Nco*I sites to the 3′ and 5′ ends of these fragments, respectively (Fig. 1). CPCR amplification yielded the expected 400-bp promoter (*P*) and 200-bp terminator (*T*) fragments shown in Fig. 2. These fragments were removed from the gel matrix, digested with *Nco*I, and ligated together to obtain the combined 600-bp fragment (*PT* in Fig. 2). This 600-bp fragment was amplified a second time using oligo primers JLS-81 and -84 (Fig. 1) to increase its concentration. The double-amplification 600-bp fragments were digested with *Hind*III, polyacrylamide gel purified, and ligated into *Hind*III-cut pUC18. After transforming *E. coli* cells five clones were selected for nt sequencing to check for ligation and/or PCR-generated mutations. This analysis showed that these clones contained identical inserts and no nt sequence mutations. This nt sequence analysis (a total of about 3000 bp from these five clones) was encouraging as it indicates that

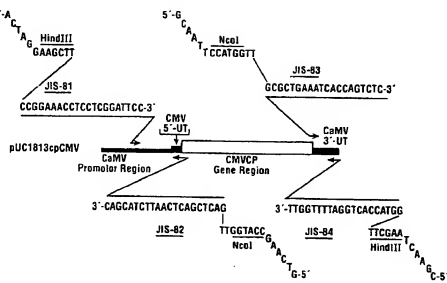

Fig. 1. Strategy used to CPCR-engineer the plant expression vector pUC18cpexp. The gene structure of clone pUC1813cpCMV19 (used for expressing the CMV *cp* gene; Quemada et al., 1991), the location and sequence of the four oligo primers used for CPCR are shown. The restriction site synthesized within each oligo primer, internal *Nco*I and flanking *Hind*III sites are also indicated (overlined or underlined). Each oligo primer contains an additional six nt adjacent to the restriction enzyme site (written on the diagonal) to aid restriction enzyme binding and DNA cleavage. Shared nt sequences between the oligo primers and the amplified regions are denoted by a line above or below the oligo sequence.

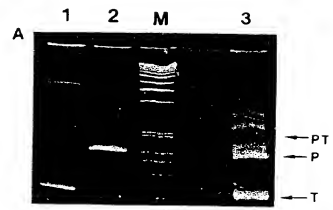

Fig 2. CPCR amplification results for the promoter and terminator regions used to construct pUC18cpexp. The amplification of 200-bp CaMV terminator (lane 1) and 400-bp CaMV-CMV promoter (lane 2) regions are shown. Lane 3 shows the results of ligating these *Nco*I-cut 200-bp and 400-bp fragments to obtain the 600-bp fragment (labeled *PT*); many fragments did not ligate as indicated by the intensity of the 200-bp terminator fragment (labeled *T*) and 400-bp promoter fragment (labeled *P*). Methods: CPCR amplification reactions contained 50 ng of each primer and 100 ng of template (pUC1818cpCMV19) in a total reaction volume of 100 μl. Buffer and nt concentrations are as described by the vendor, Perkin-Elmer-Cetus. Reaction mixtures were subjected to 30 cycles of amplification using a denaturation step of 94°C (2 min), annealing step at 50°C (1 min), and extension step of 74°C (3 min). DNA amplification cycles were controlled by a programmable cyclic reactor obtained from Ericomp, Inc. After amplification, samples were extracted once with a phenol–chloroform–isoamyl alcohol mixture (1:1:0.04), once with a chloroform–isoamyl alcohol mixture (1:0.02), precipitated with ethanol, and then resuspended in 30 μl of H_2O. Samples were then subjected to digest by *Nco*I followed by electrophoreses through a 7% polyacrylamide–15% glycerol gel. CPCR amplified bands were visualized by ethidium bromide staining and the DNA fragments were removed from the gel matrix by electroelution (Maniatis et al., 1982). These fragments were combined in equal concentration (about 0.5 μg each), then 10 units of T4 ligase (Collaborative Research, Inc.) were added and the volume adjusted to a total of 10 μl. The ligation reaction was incubated at 14°C for 18 h, after which the ligase was inactivated by heating to 65°C for ten min, then loaded on a polyacrylamide-glycerol gel. Ligation products show the presence of the expected 600-bp promoter and terminator fragment (labeled *PT*) and other products that resulted from the other possible ligation events, i.e., the 400-bp fragment (ligation of two terminator fragments) and the 800-bp (ligation of two promoter fragments).

PCR-generated errors did not occur at a frequency that would render the CPCR engineering strategy useless. The nt sequence of the CaMV-CMV regulatory regions of the CPCR-engineered expression cassette pUC18cpexp are included in Fig. 3.

(b) Cloning PRV, WMVII, and ZYMV *cp* genes into pUC18cpexp

CPCR engineering strategies were designed for the cloning of the PRV, WMVII and ZYMV *cp* genes into pUC18cpexp. These strategies were similar and straightforward as CPCR amplifications were used to add the necessary *Nco*I sites flanking each *cp* gene. The goal was to obtain a translational fusion that would upon expression produce a CP as identical as possible to the respective native CP. Some modifications were unavoidable due to the addition of an N-terminal methionine residue and because the *cp* translation fusion codon selection required conserving the *Nco*I recognition site (CCATGG). Thus the translation start codon was generally fused to the first *cp* codon beginning with a G nt. In addition, since it appeared that the preferred N-terminal aa for potyvirus CP might be Ser, Gly, or Ala (Dougherty et al., 1989), these codons were preferentially selected. Addition of the *Nco*I site to the 3′ end of these *cp* genes was much simpler because these linkages only involved transcriptional fusions.

Oligo primers designed for CPCR amplification of the PRV *cp* gene fused the translation start codon to the fifth aa (Ala) that follows the expected Glu-Ser peptide cleavage site (Quemada et al., 1990a). The 3′-*Nco*I site was added 50 bp downstream from the PRV translation stop codon (Fig. 3). CPCR amplification of the plasmid pPRV117 (Quemada et al., 1990a) yielded a 900-bp fragment that was digested with *Nco*I and then cloned into the *Nco*I-digested vector pUC18cpexp. Two selected clones were completely sequenced and no mutations were found (a total of 1800 bp were checked). Fig. 3 shows the nt sequence of CaMV-CMV regulatory elements of pUC18cpexp and its junctions with the PRV *cp* gene in the clone referred to as pUC18cpPRV.

Similar CPCR amplifications were used for the engineering of the *cp* genes from WMVII and ZYMV (data not shown). Oligo primers designed for CPCR amplification of the WMVII *cp* gene fused the translation start codon to the second aa codon, Gly, of the *cp* gene (Quemada et al., 1990b). CPCR amplification of the clone pWMVII-3.2 (Quemada et al., 1990b) yielded a 900-bp fragment that was cloned into pUC18cpexp and the clone designated pUC18cpWMVII was isolated and sequenced (data not shown), no PCR and/or ligation-generated mutations were found. CPCR amplification of the ZYMV *cp* gene fused the translation start site to the second aa, Gly, of the *cp* gene (Quemada et al., 1990b). CPCR amplification of clone pZYMV-15 (Quemada et al., 1990b) yielded a 900-bp fragment that was cloned into pUC18cpexp and the clone referred to as pUC18cpZYMV was selected and sequenced. This sequence revealed a cloning artifact at the *Nco*I site near the translation start site, it contained only one C nt instead of the expected CC nt (data not shown). The nt sequence of the remaining ZYMV *cp* gene revealed no additional errors.

(c) Conclusions

PCR amplification is a powerful method for the modification of DNAs and it was used here for the construction of

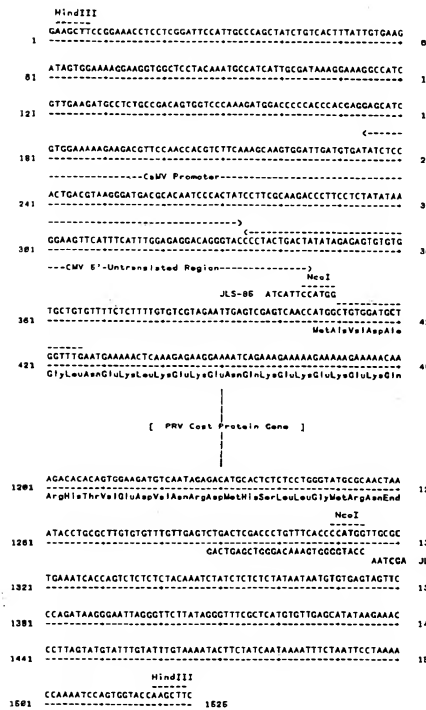

Fig. 3. The nt sequence of CPCR-engineered expression vector pUC18cpPRV. The nt sequence was obtained using double-stranded DNA (Zagursky et al., 1985) and the dideoxy sequence enzymatic sequencing method (Sequenase™2, U.S. Biochemicals). The DNA sequencing gel system used has been described by Slighton et al. (1991) and employs 1-meter gels; sequence readings in the range of 800 bp from the oligo primer are routinely obtained. The location of the CaMV promoter and terminator and CMV 5'-untranslated regions is shown. Also shown are the nt sequence and location of the oligo primers used for CPCR-amplification of the PRV *cp* gene. The location of the *Nco*I sites used for cloning the CPCR amplified PRV *cp* gene into pUC18cpexp is shown. Shared sequences between the oligo primers and PRV *cp* gene are shown as dashed lines above the sequence line and as the complementary nt sequence below the sequence line. The complete PRV *cp* gene sequence is not shown as its sequence has previously been described (Quemada et al., 1990a).

a CaMV-based plant-expression cassette. The major strength of this method is that it allows complex DNA modification to be done by oligo primer synthesis followed by PCR amplification and straightforward ligation events. A major weakness of the strategy is the concern that *Taq* polymerase lacks fidelity. However, analysis of the fidelity

of *Taq* polymerase finds it to be similar to that found for other polymerases, T7 and DNA polymerase I (Keohavong and Thilly, 1989). In addition, the probability of obtaining PCR-generated errors in the examples presented here are reduced because the initial template is cloned and its concentration is not limiting, as is the case when amplifying rare mRNAs or single-copy genomic DNAs. This is supported by nt sequence analyses of the CPCR-engineered clones described here (a total of 8400 bp) since no PCR-generated errors were found. However, it was surprising to find an error that resulted from a ligation event; thus, it is recommended that CPCR-engineered DNAs be sequenced to ensure accuracy.

The constitutive plant expression vector pUC18cpexp can be effectively used for the cloning of almost any coding region provided that it is modified by CPCR amplification. However, coding regions that contain multiple *Nco*I sites will be problematic and may require the selection of a different restriction enzyme site to be placed at or near the translation start codon. Each of the potyvirus *cp* expression vectors described here has been transferred into plant tissues and found to be functional (S. Namba, K. Ling, C. Gonsalves, J.L.S. and D. Gonsalves, manuscript in preparation).

ACKNOWLEDGEMENTS

I wish to thank Nicole Hatzenbuhler and Mary Shea for the synthesis of oligo primers and Leang Sieu for nt sequencing.

REFERENCES

Dougherty, W.G. and Carrington, J.C.: Expression and function of potyviral gene products. Annu. Rev. Phytopathol. 26 (1988) 123–143.

Dougherty, W.G., Cary, S.M. and Parks, T.D.: Molecular genetic analysis of a plant virus polyprotein cleavage site: a model. Virology 171 (1989) 356–364.

Gallie, D.R., Sleat, D.E., Watts, J.W., Turner, P.C. and Wilson, T.M.A.: The 5′-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts in vitro and in vivo. Nucleic Acids Res. 15 (1987) 3257–3273.

Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J.: PCR Protocols: A Guide to Methods and Applications. Academic Press, New York, 1989.

Jobling, S.A. and Gehrke, L.: Enhanced translation of chimaeric messenger RNAs containing a plant viral untranslated leader sequence. Nature 325 (1987) 622–625.

Keohavong, P. and Thilly, W.G.: Fidelity of DNA polymerases in DNA amplification. Proc. Natl. Acad. Sci. USA 86 (1989) 9253–9257.

Lutcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F. and Scheele, G.A.: Selection of AUG initiation codons differs in plants and animals. EMBO J. 6 (1987) 43–48.

MacFerrin, K.D., Terranoma, M.P., Schreiber, S.L. and Verdine, G.L.: Overproduction and dissection of proteins by the expression-cassette polymerase chain reaction. Proc. Natl. Acad. Sci. USA 87 (1990) 1937–1941.

Maniatis, T., Fritsch, E. and Sambrook, J.: Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.

Pietrzak, M., Shillito, R.D., Hohn, T. and Potrykus, I.: Expression in plants of two bacterial antibiotic resistance genes after protoplast transformation with a new plant expression vector. Nucleic Acids Res. 14 (1986) 5857–5868.

Quemada, H., Kearney, C., Gonsalves, D. and Slightom, J.L.: Nucleotide sequences of the coat protein genes and flanking regions of cucumber mosaic virus strains C and WL RNA 3. J. Gen. Virol. 70 (1989) 1065–1073.

Quemada, H., L'Hostis, B., Gonsalves, D., Reardon, I., Heinrickson, R., Hiebert, E.L., Sieu, L.C. and Slightom, J.L.: The nucleotide sequences of the 3′ terminal regions of papaya ringspot virus strains W and P. J. Gen. Virol. 71 (1990a) 203–210.

Quemada, H., Sieu, L.C., Siemieniak, D.R., Gonsalves, D. and Slightom, J.L.: Watermelon mosaic virus II and Zucchini yellow mosaic virus: cloning of 3′-terminal regions, nucleotide sequences, and phylogenic comparisons. J. Gen. Virol. 71 (1990b) 1451–1460.

Quemada, H.D., Rodriguez, R., Gonsalves, D. and Slightom, J.L.: CMV-C coat protein gene expression in tobacco: protection against infections by CMV strains transmitted mechanically or by aphids. Phytopathology (1991) in press.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A.: Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239 (1988) 487–491.

Slightom, J.L., Siemieniak, D.R. and Sieu, L.C.: DNA sequencing: strategy and methods to directly sequence large DNA molecules. In: Miyamoto, M.M. and Cracraft, J. (Eds.), Phylogenetic Analysis of DNA. Oxford University Press, Oxford, U.K., 1991, in press.

Zagursky, R.J., Baumeister, K., Lomax, N. and Berman, M.L.: Rapid and easy sequencing of large double-stranded and supercoiled plasmid DNA. Gene Analyt. Tech. 2 (1985) 89–94.